DERIVATIONS WITH ENGEL CONDITIONS
ON MULTILINEAR POLYNOMIALS

PJEK–HWEE LEE AND TSIU–KWEN LEE

(Communicated by Ken Goodearl)

Abstract. Let $R$ be a prime algebra over a commutative ring $K$ with unity and let $f(x_1, \ldots, x_n)$ be a multilinear polynomial over $K$. Suppose that $d$ is a nonzero derivation on $R$ such that for all $r_1, \ldots, r_n$ in some nonzero ideal $I$ of $R$, $\left[ d(f(r_1, \ldots, r_n)), f(r_1, \ldots, r_n) \right]_k = 0$ with $k$ fixed. Then $f(x_1, \ldots, x_n)$ is central-valued on $R$ except when $\text{char } R = 2$ and $R$ satisfies the standard identity $s_4$ in 4 variables.

Throughout this note $K$ will denote a commutative ring with unity and $R$ will denote a prime $K$–algebra with center $Z$. By $d$ we always mean a nonzero derivation on $R$. For $x, y \in R$, set $[x, y]_0 = x$, $[x, y]_1 = [x, y] = xy - yx$ and, for $k > 1$, $[x, y]_k = \left[ [x, y]_{k-1}, y \right]$. A well–known result proved by Posner [10] states that $R$ must be commutative if $[d(x), x] \in Z$ for all $x \in R$. In [7], the authors generalized Posner’s theorem by showing that a Lie ideal $L$ of $R$ must be contained in $Z$ if $\text{char } R \neq 2$ and $[d(x), x] \in Z$ for all $x \in L$. As to the case when $\text{char } R = 2$, Lanski [5] obtained the same conclusion except when $R$ satisfies the standard identity $s_4$ in 4 variables. On the other hand, Vukman [11] showed that $R$ is commutative if $\text{char } R \neq 2$ and $[d(x), x]_2 = 0$ for all $x \in R$, or if $\text{char } R \neq 2, 3$ and $[d(x), x]_2 \in Z$ for all $x \in R$.

In a recent paper [6], a full generalization of these results was proved by Lanski. He showed that a Lie ideal $L$ of $R$ is in $Z$ if for some fixed $k > 0$, $[d(x), x]_k = 0$ for all $x \in L$, unless $\text{char } R = 2$ and $R$ satisfies $s_4$. Note that a noncentral Lie ideal of $R$ contains all the commutators $[x, y]$ for $x, y$ in some nonzero ideal of $R$ except when $\text{char } R = 2$ and $R$ satisfies $s_4$. It is natural to consider the situation when $[d(x), x]_k = 0$ for all commutators $x = [x_1, x_2]$, or more generally, when $[d(x), x]_k = 0$ for all $x = f(x_1, \ldots, x_n)$ where $f(x_1, \ldots, x_n)$ is a multilinear polynomial over $K$. In the present paper, we shall extend Lanski’s theorem by imposing the condition $\left[ d(f(x_1, \ldots, x_n)), f(x_1, \ldots, x_n) \right]_k = 0$ on some nonzero ideal of $R$. First we dispose of the simplest case when $R$ is the matrix ring $M_m(F)$ over a field $F$ and $d$ is an inner derivation on $R$.

\[ \text{Received by the editors November 4, 1994 and, in revised form, March 1, 1995.} \]

\[ 1991 \text{ Mathematics Subject Classification.} \] Primary 16W25; Secondary 16N60, 16R50, 16U80.

\[ \text{Key words and phrases.} \] Multilinear polynomial, derivation, generalized polynomial identity.

\[ \text{©1996 American Mathematical Society} \]

\[ 2625 \]
Proposition. Let $F$ be a field and $R = M_m(F)$, the $m \times m$ matrix algebra over $F$. Suppose that $a \in R$ and that $f(x_1, \ldots, x_n)$ is a multilinear polynomial over $F$ such that $[a, f(x_1, \ldots, x_n)]_k = 0$ for all $x_i \in R$, where $k > 0$ is a fixed integer. Then either $a \in Z$ or $f(X_1, \ldots, X_n)$ is central–valued on $R$ unless $\text{char } F = 2$ and $m = 2$.

Proof. If $m = 1$, there is nothing to prove, so we assume that $m \geq 2$. We assume further that either $m \neq 2$ or $\text{char } F \neq 2$ and proceed to show that $a \in Z$ if $f(X_1, \ldots, X_n)$ is not central–valued on $R$. By [8, Lemmas 2 and 9], there exists a sequence of matrices $r = (r_1, \ldots, r_n)$ in $R$ such that $f(r) = f(r_1, \ldots, r_n) = \sum_{i=1}^m a_i e_{ii}$ is a noncentral diagonal matrix, say, $a_1 \neq a_2$. Let $\varphi$ be the inner automorphism on $R$ defined by $x^\varphi = (1 + e_{21})x(1 - e_{21})$. Write $D = f(r)$; then $E = f(r^\varphi) = f(r_1^\varphi, \ldots, r_n^\varphi) = D^\varphi = D + (a_1 - a_2)e_{21}$. By assumption, we have $[a, E]_k = 0$, that is, $\sum_{i=0}^k (-1)^i \binom{k}{i} E^i a E^{k-i} = 0$. Multiplying $e_{22}$ from both sides, we have $(a_2 - a_1)e_{22} = 0$ by using $E^i = D^i + (a_1 - a_2)e_{21}$, $e_{22} E^i = a_2^i e_{22} - e_{21} + a_1^i e_{21}$ and $E^i e_{22} = a_2^i e_{22}$. Thus, $a_2^i e_{22} = 0$ and so the $(1, 2)$–entry of $a$ is $0$. For $s \neq t$, let $\sigma$ be a permutation in the symmetric group $S_m$ such that $\sigma(1) = s$ and $\sigma(2) = t$. Let $\psi$ be the automorphism on $R$ defined by $x^\psi = \left( \sum_{i,j} \xi_{ij} e_{ij} \right)^\psi = \sum_{i,j} \xi_{ij} e_{\sigma(i), \sigma(j)}$. Then $f(r^\psi) = D^\psi$ is a diagonal matrix with $(s, s)$– and $(t, t)$–entries distinct. Replacing $D$ by $D^\psi$ and proceeding as before, we have $a_{st} = 0$ for $s \neq t$. In other words, $a$ is a diagonal matrix. For any $F$–automorphism $\theta$ of $R$, $a^\theta$ enjoys the same property as $a$ does, namely, $[a^\theta, f(x_1, \ldots, x_n)]_k = 0$ for all $x_i \in R$. Hence, $a^\theta$ must be diagonal. Write $a = \sum_{i=1}^m a_{ii} e_{ii}$; then for $s \neq t$, we have $(1 + e_{ts}) a (1 - e_{ts}) = \sum_{i=1}^m a_{ii} e_{ii} + (a_{ss} - a_{tt}) e_{ts}$ diagonal. Hence, $a_{ss} = a_{tt}$ and so $a$ is a scalar matrix.

We are now ready to prove the main

Theorem. Let $R$ be a prime $K$–algebra and $f(x_1, \ldots, x_n)$ a multilinear polynomial over $K$. Suppose that $d$ is a nonzero derivation on $R$ such that

$$[#(f(x_1, \ldots, x_n), f(x_1, \ldots, x_n))]_k = 0$$

for all $x_i$ in some nonzero ideal $I$ of $R$, where $k > 0$ is fixed. Then $f(X_1, \ldots, X_n)$ is central–valued on $R$ except when $\text{char } R = 2$ and $R$ satisfies $S_4$.

Proof. Assume first that $d$ is $Q$–inner, that is, $d(x) = [a, x]$ for all $x \in R$, where $a$ is a noncentral element in the symmetric quotient ring $Q$ of $R$ [4]. Then $[a, f(x_1, \ldots, x_n)]_{k+1} = 0$ for all $x_i \in I$.

By a theorem due to Chuang [1, Theorem 2], this generalized polynomial identity $[a, f(X_1, \ldots, X_n)]_{k+1}$ is also satisfied by $Q$. In case the center $C$ of $Q$ is infinite, we
have \([a, f(x_1, \ldots, x_n)]_{k+1} = 0\) for all \(x_i \in Q \otimes_C \overline{C}\) where \(\overline{C}\) is the algebraic closure of \(C\). Since both \(Q\) and \(Q \otimes_C \overline{C}\) are prime and centrally closed [2, Theorems 2.5 and 3.5], we may replace \(R\) by \(Q\) or \(Q \otimes_C \overline{C}\) according as \(C\) is finite or infinite. Thus we may assume that \(R\) is centrally closed over \(C\) which is either finite or algebraically closed and \([a, f(x_1, \ldots, x_n)]_{k+1} = 0\) for all \(x_i \in R\).

By Martindale’s theorem [9], \(R\) is then a primitive ring having a nonzero socle \(H\) with \(C\) as the associated division ring. In light of Jacobson’s theorem [3, p. 75], \(R\) is isomorphic to a dense ring of linear transformations on some vector space \(V\) over \(C\), and \(H\) consists of the linear transformations in \(R\) of finite rank. Assume first that \(V\) is finite–dimensional over \(C\). Then the density of \(R\) on \(C^V\) implies that \(R \cong M_m(C)\) where \(m = \dim_C V\). By the preceding proposition, \(f(X_1, \ldots, X_n)\) is central–valued on \(R\) unless \(\text{char } R = 2\) and \(n = 2\). Assume next that \(V\) is infinite–dimensional over \(C\). For any \(e = e^3 \in H\), we have \(eRe \cong M_m(C)\) with \(m = \dim_C Ve\). Since \(R\) satisfies \(e[a, f(ex_1, \ldots, ex_n)]_{k+1} = 0\), \(eRe\) satisfies \([eae, f(X_1, \ldots, x_n)]_{k+1} = 0\). As we have seen above, \(eae\) must be central in \(eRe\) if \(m \geq 3\) and if \(f(X_1, \ldots, X_n)\) is not central–valued on \(eRe\). Since \(a \notin C\), \(a\) does not centralize the nonzero ideal \(H\) of \(R\), so \(ah_0 \neq h_0a\) for some \(h_0 \in H\). Also, \(f(X_1, \ldots, X_n)\) is not central–valued on \(H\), for otherwise \(R\) would satisfy the polynomial identity \([f(X_1, \ldots, X_n), X_{n+1}]\) contrary to the infinite–dimensionality of \(C^V\). So, \([f(h_1, \ldots, h_n), h_{n+1}] \neq 0\) for some \(h_1, \ldots, h_n, h_{n+1} \in H\). By Litoff’s theorem [5, p. 280], there is an idempotent \(e \in H\) so that \(h_0, h_0a, ah_0, h_1, \ldots, h_n, h_{n+1}\) are all in \(eRe\). Since \(\dim_C V\) is infinite, we may choose \(e\) so that \(m = \dim_C Ve \geq 3\). Then \(eae\) is central in \(eRe\) because \(f(X_1, \ldots, X_n)\) is not central–valued on \(eRe\). Thus

\[
ah_0 = eah_0 = eaeh_0 = h_0eae = h_0ae = h_0a,
\]

a contradiction. Hence \(f(X_1, \ldots, X_n)\) must be central–valued on \(R\) except when \(\text{char } R = 2\) and \(R\) satisfies \(s_4\).

Now assume that \(d\) is not \(Q\)-inner. Recall that \(d\) can be extended uniquely to a derivation \(\overline{d}\) on \(Q\) [4]. We denote by \(f^d(X_1, \ldots, X_n)\) the polynomial obtained from \(f(X_1, \ldots, X_n)\) by replacing each coefficient \(a\) with \(\overline{d} (a \cdot 1)\). Since

\[
\left[ d(f(x_1, \ldots, x_n)), f(x_1, \ldots, x_n) \right]_k \\
= \left[ f^d(x_1, \ldots, x_n) + d(f(x_1), x_2, \ldots, x_n) \\
+ \cdots + f(x_1, \ldots, x_{n-1}, d(x_n)), f(x_1, \ldots, x_n) \right]_k = 0,
\]

we have

\[
f^d(x_1, \ldots, x_n) + f(y_1, x_2, \ldots, x_n) \\
+ \cdots + f(x_1, \ldots, x_{n-1}, y_n), f(x_1, \ldots, x_n) \right]_k = 0
\]

for all \(x_1, \ldots, x_n, y_1, \ldots, y_n \in R\) by Kharchenko’s theorem [4]. In particular, \([f^d(x_1, \ldots, x_n), f(x_1, \ldots, x_n)]_k = 0\) and \([f(y_1, x_2, \ldots, x_n), f(x_1, \ldots, x_n)]_k = 0\) for all \(x_1, \ldots, x_n, y_1 \in R\). Applying \(d\) to \([f(y_1, x_2, \ldots, x_n), f(x_1, \ldots, x_n)]_{2k} = 0\),
we have
\[
\left[ f^d(y_1, x_2, \ldots, x_n) + f(d(y_1), x_2, \ldots, x_n) \right]
\]
\[+ \sum_{i=2}^{n} f(y_1, x_2, \ldots, d(x_i), \ldots, x_n), f(x_1, \ldots, x_n) \right]_{2k}
\]
\[+ \sum_{j=0}^{2k-1} \left[ \left[ f(y_1, x_2, \ldots, x_n), f(x_1, \ldots, x_n) \right]_j, d(f(x_1, \ldots, x_n)) \right] \]
\[f(x_1, \ldots, x_n)_{2k-j-1} = 0.
\]
The second sum vanishes since, for each \(j\),
\[
\left[ \left[ f(y_1, x_2, \ldots, x_n), f(x_1, \ldots, x_n) \right]_j, d(f(x_1, \ldots, x_n)) \right]_{2k-j-1}
\]
\[= \sum_{i=0}^{2k-j-1} \binom{2k-j-1}{i} \left[ f(y_1, x_2, \ldots, x_n), f(x_1, \ldots, x_n) \right]_{j+i},
\]
\[d(f(x_1, \ldots, x_n)) = 0,
\]
where either \(j + i \geq k\) or \(2k - j - i - 1 \geq k\). Thus we have
\[
\left[ f^d(y_1, x_2, \ldots, x_n) + f(d(y_1), x_2, \ldots, x_n) \right]
\]
\[+ \sum_{i=2}^{n} f(y_1, x_2, \ldots, d(x_i), \ldots, x_n), f(x_1, \ldots, x_n) \right]_{2k} = 0.
\]
Again by Kharchenko’s theorem, we have
\[
\left[ f^d(y_1, x_2, \ldots, x_n), f(x_1, \ldots, x_n) \right]_{2k}
\]
\[+ f(y_1, y_2, x_3, \ldots, x_n) + \cdots + f(y_1, x_2, \ldots, y_n), f(x_1, \ldots, x_n) \right]_{2k} = 0
\]
for all \(x_1, \ldots, x_n, y_1, \ldots, y_n, z \in R\). In particular,
\[
\left[ f^d(y_1, x_2, \ldots, x_n), f(x_1, \ldots, x_n) \right]_{2k} = 0
\]
and
\[
\left[ f(y_1, y_2, x_3, \ldots, x_n), f(x_1, \ldots, x_n) \right]_{2k} = 0
\]
for all \(x_1, \ldots, x_n, y_1, y_2 \in R\). Continuing this process, we will finally get
\[
\left[ f(y_1, \ldots, y_n), f(x_1, \ldots, x_n) \right]_{nk} = 0
\]
for all \(x_1, \ldots, x_n, y_1, \ldots, y_n \in R\). In light of the inner case, we have \(f(y_1, \ldots, y_n) \in Z\) for all \(y_1, \ldots, y_n \in R\), that is, \(f(X_1, \ldots, X_n)\) is central-valued on \(R\), except when \(\text{char } R = 2\) and \(R\) satisfies \(s_4\).

\begin{thebibliography}{3}
\end{thebibliography}

Department of Mathematics, National Taiwan University, Taipei, Taiwan
E-mail address: phlee@math.ntu.edu.tw

Department of Mathematics, National Taiwan University, Taipei, Taiwan
E-mail address: tklee@math.ntu.edu.tw